

## Article

# Quantitative Ultrasound-Based Precision Diagnosis of Papillary, Follicular, and Medullary Thyroid Carcinomas Using Morphological, Structural, and Textural Features

Hanna Piotrkowska Wróblewska <sup>1,\*</sup>, Piotr Karwat <sup>1</sup>, Agnieszka Żyłka <sup>2</sup>, Katarzyna Dobruch Sobczak <sup>3</sup>, Marek Dedecjus <sup>2</sup> and Jerzy Litniewski <sup>1</sup>

<sup>1</sup> Ultrasound Department, Institute of Fundamental Technological Research, Polish Academy of Sciences, 02-106 Warsaw, Poland

<sup>2</sup> Department of Endocrine Oncology and Nuclear Medicine, Maria Skłodowska-Curie National Research Institute of Oncology, 02-781 Warsaw, Poland

<sup>3</sup> Radiology Department II, Maria Skłodowska-Curie National Research Institute of Oncology, 02-034 Warsaw, Poland; kdsobczak@gmail.com

\* Correspondence: hpiotrk@ippt.pan.pl

## Simple Summary

Thyroid cancer includes several types that differ in how they grow and how they should be treated. Although ultrasound is widely used to examine thyroid nodules, it can be difficult to determine which type of cancer is present using standard imaging alone. In this study, we applied a computer-based method to automatically measure and analyze ultrasound features of thyroid tumors. By using machine learning techniques, we distinguished between three common types of thyroid cancer: papillary, follicular, and medullary. We found that certain features, such as tumor shape, brightness, and internal structure, were helpful in identifying the cancer subtype. This approach could support doctors in making more accurate diagnoses, reduce unnecessary procedures such as biopsies, and guide more personalized treatment decisions.



Academic Editor: Michael Kreißl

Received: 27 July 2025

Revised: 22 August 2025

Accepted: 23 August 2025

Published: 24 August 2025

**Citation:** Piotrkowska Wróblewska, H.; Karwat, P.; Żyłka, A.; Dobruch Sobczak, K.; Dedecjus, M.; Litniewski, J. Quantitative Ultrasound-Based Precision Diagnosis of Papillary, Follicular, and Medullary Thyroid Carcinomas Using Morphological, Structural, and Textural Features. *Cancers* **2025**, *17*, 2761. <https://doi.org/10.3390/cancers17172761>

**Copyright:** © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## Abstract

**Background/Objectives:** Thyroid cancer encompasses distinct histological subtypes with varying biological behavior and treatment implications. Accurate preoperative subtype differentiation remains challenging. Although ultrasound (US) is widely used for thyroid nodule evaluation, qualitative assessment alone is often insufficient to distinguish between papillary (PTC), follicular (FTC), and medullary thyroid carcinoma (MTC). **Methods:** A retrospective analysis was performed on patients with histologically confirmed PTC, FTC, or MTC. A total of 224 standardized B-mode ultrasound images were analyzed. A set of fully quantitative features was extracted, including morphological characteristics (aspect ratio and perimeter-to-area ratio), internal echotexture (echogenicity and local entropy), boundary sharpness (gradient measures and KL divergence), and structural components (calcifications and cystic areas). Feature extraction was conducted using semi-automatic algorithms implemented in MATLAB. Statistical differences were assessed using the Kruskal–Wallis and Dunn–Šidák tests. A Random Forest classifier was trained and evaluated to determine the discriminatory performance of individual and combined features. **Results:** Significant differences ( $p < 0.05$ ) were found among subtypes for key features such as perimeter-to-area ratio, normalized echogenicity, and calcification pattern. The full-feature Random Forest model achieved an overall classification accuracy of 89.3%, with F1-scores of 93.4% for PTC, 85.7% for MTC, and 69.1% for FTC. A reduced model using the top 10 features yielded an even higher accuracy of 91.8%, confirming the robustness

and clinical relevance of the selected parameters. **Conclusions:** Subtype classification of thyroid cancer was effectively performed using quantitative ultrasound features and machine learning. The results suggest that biologically interpretable image-derived metrics may assist in preoperative decision-making and potentially reduce the reliance on invasive diagnostic procedures.

**Keywords:** thyroid cancer; ultrasound imaging; quantitative analysis; machine learning; papillary thyroid carcinoma; follicular thyroid carcinoma; medullary thyroid carcinoma

## 1. Introduction

Thyroid cancer (TC) is the most common malignancy of the endocrine system, and its global incidence has steadily increased in recent decades. According to GLOBOCAN 2020 data, more than 586,000 new TC cases are diagnosed each year, with a strong predominance in women [1]. In Poland alone, over 4000 cases were reported in 2021, confirming the significance of this disease at the population level [2]. Although this trend is partly due to improved access to high-resolution ultrasound (US) and widespread screening, a true increase in incidence is also suspected [3,4].

Thyroid malignancies encompass diverse histopathological subtypes with distinct biological behavior, prognoses, and treatment strategies. Papillary thyroid carcinoma (PTC) is the most common subtype (80–85%) and is generally associated with excellent prognosis and indolent progression [4–6]. Follicular thyroid carcinoma (FTC), comprising 10–15% of cases, may follow a more aggressive course, especially in the presence of vascular invasion or distant metastases [4–6]. Medullary thyroid carcinoma (MTC), arising from parafollicular C cells, accounts for 1–2% of thyroid cancers and often presents as part of inherited syndromes such as MEN 2A and 2B. Due to its neuroendocrine origin, MTC requires additional biochemical testing (e.g., serum calcitonin), genetic screening, and a different surgical approach [6–8].

Accurate preoperative differentiation between these subtypes—particularly between PTC, FTC, and MTC—is essential for treatment planning. Ultrasound remains the primary diagnostic tool for evaluating thyroid nodules, as it is safe, widely accessible, and highly sensitive for solid lesions. Multiparametric ultrasound further enables detailed assessment of lesion morphology, facilitates initial risk stratification, and supports biopsy qualification. Various risk stratification systems, including ACR TI-RADS, EU-TIRADS, and EU-TIRADS-PL, have been developed to standardize nodule descriptions and reduce unnecessary fine-needle aspiration biopsies (FNABs) [9–11]. However, these systems only assess malignancy risk without providing insight into the histopathological subtype.

The ability to distinguish PTC, FTC, and MTC based on sonographic features may provide critical support in clinical decision-making [12–14]. Although several studies have described typical ultrasound features of individual subtypes, comprehensive comparative analyses directly contrasting PTC, FTC, and MTC remain scarce [15–23]. Table 1 summarizes the most frequently reported ultrasound features for these three main subtypes, highlighting differences in echogenicity, margins, calcifications, vascularity, and stiffness [12–23].

Medullary thyroid carcinoma (MTC) can often mimic benign lesions on ultrasound. It typically presents as a larger, solid, hypoechoic nodule with smooth margins, rich internal vascularity, and coarse (macro-) calcifications. Compared to papillary thyroid carcinoma (PTC), MTC less frequently displays irregular margins or microcalcifications and is more likely to exhibit increased intranodular blood flow. The ultrasound appearance of MTC can

be difficult to distinguish from that of benign nodules, particularly when smooth margins are present [13,24,25].

**Table 1.** Most commonly reported ultrasound features of the three main thyroid cancer subtypes.

Ultrasound Feature	Papillary Thyroid Carcinoma (PTC)	Follicular Thyroid Carcinoma (FTC)	Medullary Thyroid Carcinoma (MTC)
Echogenicity	Hypoechoic, sometimes heterogeneous [19–22]	Iso- or hypoechoic [18,20–22]	Hypoechoic [13,19,23]
Margins	Irregular, ill-defined [19–22]	Regular or irregular if invasive [18,20–22]	Smooth, well-defined (sometimes ill-defined) [13,19,23]
Calcifications	Microcalcifications (psammoma bodies) [19–21]	Micro- and macrocalcifications, often peripheral (“eggshell”) [18,20,21]	Micro- and macrocalcifications (amyloid deposits, shadowing) [13,19,23]
Internal structure	Solid, possibly heterogeneous [19–22]	Solid, heterogeneous [18,20,21]	Solid, possibly homogeneous [13,19]
Shape (aspect ratio)	“Taller-than-wide” (common) [18–21]	Oval or irregular [18,20,21]	Variable, round or oval [13,19]
Vascularity (Doppler)	Often increased, chaotic internal pattern [18,19]	Moderate, mixed pattern [17,18]	Increased, central and peripheral [13,17,19]
Elastography	High stiffness [19]	Variable, often intermediate [18]	High stiffness [18,23]
Presence of capsule	Absent or interrupted capsule [18,21]	Often infiltrated, extracapsular extension [18,21]	Absent [18]
Cystic component	Rare, usually <10% of volume [18,19]	Rare [18]	Often present, especially in larger lesions [13]
Lymph node metastases	Common at diagnosis [18,19]	Less common [13,19]	Common [13,19]

In contrast, PTC is more commonly associated with features suggestive of malignancy, such as irregular margins, microcalcifications, hypoechogenicity, and a “taller-than-wide” shape. Classic PTC typically demonstrates more ultrasound features of malignancy compared to the follicular variant (FVPTC), which more frequently shows smooth margins and fewer microcalcifications. Consequently, FVPTC may be more challenging to identify based solely on ultrasound appearance, as it often lacks the typical high-risk features [22,26,27].

Despite advances in ultrasound imaging techniques, differentiating thyroid cancer subtypes based solely on B-mode imaging remains a diagnostic challenge due to considerable overlap in sonographic features. In recent years, there has been growing interest in the application of quantitative image analysis and machine learning algorithms in the evaluation of thyroid nodules [28,29]. However, the majority of existing studies have focused on binary classification (benign vs. malignant), without accounting for histological differentiation.

The aim of this study was to evaluate whether automatically extracted multiparametric ultrasound features can be used to distinguish between the three major histological subtypes of thyroid cancer: papillary (PTC), follicular (FTC), and medullary (MTC). The objective was to assess the diagnostic value of quantitative imaging parameters in supporting accurate subtype classification and informing personalized therapeutic strategies.

The novelty of our work lies in moving beyond the binary paradigm of thyroid nodule assessment. While most prior ultrasound studies have focused on distinguishing benign from malignant lesions, direct quantitative comparisons among papillary, follicular, and medullary carcinomas remain scarce. In contrast to qualitative TI-RADS descriptors, which are subjective and prone to interobserver variability, we employed fully quantitative and standardized imaging features that are clinically meaningful. Our approach integrates multiple domains of information—morphological shape metrics, echogenicity, margin sharpness, calcification distribution, and textural parameters—into a unified classification model. Compared with previous ultrasound-based studies relying mainly on qualitative descriptors or handcrafted single-domain features [15–23], our method leverages a systematic, multiparametric, and automated analysis pipeline. Furthermore, unlike CT or MRI, which

provide complementary but less accessible diagnostic information, B-mode ultrasound is widely available, cost-effective, and safe. By enhancing its diagnostic capability through quantitative analysis, our method aims to provide a practical and transparent tool for improving preoperative subtype differentiation.

## 2. Materials and Methods

### 2.1. Study Design and Patient Cohort

The study included patients who underwent surgery between 2021 and 2022 at the Department of Oncological Endocrinology and Nuclear Medicine, Maria Skłodowska-Curie National Research Institute of Oncology in Warsaw. The initial dataset comprised 214 thyroid nodules that were evaluated by ultrasound and subjected to ultrasound-guided fine-needle aspiration biopsy (FNAB). This cohort included benign lesions, borderline tumors, and nodules with malignant potential. For the purpose of this study, only cases with postoperative histopathological confirmation of one of the three main malignant subtypes—papillary thyroid carcinoma (PTC,  $n = 90$ ), follicular thyroid carcinoma (FTC,  $n = 14$ ), or medullary thyroid carcinoma (MTC,  $n = 18$ )—were included, while benign nodules, borderline tumors, and anaplastic thyroid carcinomas were excluded. In patients with multiple nodules, the dominant lesion—defined as the largest or most suspicious on ultrasound—was selected for analysis. The final study group consisted of 122 patients (104 women and 18 men).

Clinical information was collected for all cases. The mean age of the patients was 48.3 years (range: 22–85). Lesions were located in the right lobe in 62 cases, the left lobe in 53 cases, and the isthmus in 7 cases. Nodule size was measured in three orthogonal dimensions (anteroposterior, transverse, and longitudinal), with the maximum diameter used for descriptive statistics (mean: 19.1 mm, range: 4–92 mm). According to the EU-TIRADS classification, 9 nodules were categorized as EU-TIRADS 3, 16 as EU-TIRADS 4, and 97 as EU-TIRADS 5. Cytological findings were classified according to the Bethesda System: 4 nodules were category II, 4 were category III (AUS/FLUS), 19 were category IV, 57 were category V, and 38 were category VI.

Pathological staging was determined based on the AJCC/UICC TNM classification. The majority of lesions were staged as pT1aN0 ( $n = 47$ ) and pT1bN0 ( $n = 21$ ), with smaller groups corresponding to pT1aNx ( $n = 4$ ), pT1aN1a ( $n = 2$ ), pT1aN1b ( $n = 1$ ), pT1bNx ( $n = 7$ ), pT1bN1a ( $n = 2$ ), and pT1bN1b ( $n = 6$ ). For T2 tumors, staging included pT2Nx ( $n = 5$ ), pT2N0 ( $n = 14$ ), pT2N1a ( $n = 1$ ), and pT2N1b ( $n = 1$ ). Less frequent were T3 lesions: pT3aNx ( $n = 1$ ), pT3aN0 ( $n = 5$ ), and pT3aN1b ( $n = 4$ ). Only one case was classified as pT4aN1b, while no tumors were staged as pT3b or pT4b.

Surgical treatment included either total thyroidectomy or lobectomy with isthmectomy, and in many cases it was complemented by central neck compartment lymphadenectomy. The extent of surgery was determined individually by the surgical oncologist, based on clinical presentation, ultrasound findings, and cytological evaluation.

Ultrasound images were anonymized, and analyses were conducted retrospectively using archived data. The structure of the study cohort reflected the known epidemiological predominance of PTC over other thyroid cancer subtypes [30–33].

The study was approved by the Bioethics Committee of the Maria Skłodowska-Curie National Research Institute of Oncology in Warsaw (approval number 83/2021). Written informed consent was obtained from all participants prior to inclusion in the study.

### 2.2. Image Acquisition and Preprocessing

B-mode ultrasound images of focal thyroid lesions were acquired from 122 patients with histopathologically confirmed malignant tumors. For each case, two orthogonal

images—transverse and longitudinal—were obtained, resulting in a total of 244 images for further analysis.

All examinations were performed using a Philips Epiq 5 ultrasound system equipped with a high-frequency linear transducer (eL18-4, 4–18 MHz). Images were acquired in B-mode at a central frequency of approximately 12–18 MHz, with depth settings ranging from 3.0 to 4.2 cm to fully visualize the thyroid gland and the focal lesion. The dynamic range was set to 68 dB, and overall gain was adjusted individually for each patient within the range of 40–55% to optimize image contrast. A single focal zone was positioned at the center of the lesion and, in some cases, at the lower margin of the lesion to optimize boundary visualization. Both transverse and longitudinal planes were acquired for each nodule.

Subsequent image processing was performed in the MATLAB environment (MathWorks, Natick, MA, USA). Binary masks of each lesion were generated, initially created manually by physicians and then refined semi-automatically using morphological operations and an active contour algorithm. The agreement between the initial and final mask was assessed using the Dice similarity coefficient ( $0.9637 \pm 0.013$ ), which in this context reflected the extent of correction introduced by the algorithm relative to the manual segmentation. It should be noted that, in our study, the Dice coefficient was not used to compare the result to an independent reference mask, but rather to quantify the degree of modification introduced by the algorithm relative to the initial segmentation. The spatial resolution of the images was determined based on metadata from the ultrasound device, enabling the conversion of all pixel-based measurements into millimeter (mm) units. This allowed for reliable extraction of geometric and structural features of the analyzed lesions.

All ultrasound images and their corresponding segmentation masks were anonymized.

## 2.3. Quantitative Feature Extraction

### 2.3.1. Morphological Features

- Aspect Ratio (Height-to-Width Ratio)

One of the fundamental and most commonly used morphological features of focal thyroid lesions in clinical practice is their size and shape, including the height-to-width ratio, known as the aspect ratio. For each lesion, a minimum bounding rectangle was defined in the transverse plane, in accordance with the recommendations of the American Thyroid Association and the EU-TIRADS and ACR-TIRADS systems [4,9–11].

Mathematically, the aspect ratio was defined as follows:

$$AR = \frac{H}{W} \quad (1)$$

where  $H$  denotes the anteroposterior (height) dimension of the lesion and  $W$  the transverse (width) dimension. A value greater than 1 (so-called “taller-than-wide”) indicates predominant growth along the anteroposterior axis, which is typical for infiltrative lesions and is associated with a higher risk of malignancy [34–38].

In the context of thyroid cancer subtype differentiation, this feature is particularly characteristic of papillary thyroid carcinoma (PTC), which more frequently exhibits a “taller-than-wide” shape compared to follicular (FTC) or medullary thyroid carcinoma (MTC). In contrast, FTC and MTC more often present with regular proportions, which may lead to misclassification as benign lesions [14,36,39–41].

- Shape Complexity (Perimeter-to-Area Ratio)

To obtain accurate binary masks of focal thyroid lesions, a semi-automatic refinement method was applied, based on initial manual segmentations performed by experienced clinicians. The segmentation process began with morphological opening to eliminate small

artifacts while preserving the overall integrity of the lesion shape. Subsequently, the mask contour was refined using an active contour model, which minimized the total contour energy by incorporating an internal term (promoting contour smoothness) and an external term (attracting the contour toward lesion boundaries based on image intensity gradients). The active contour algorithm was iterated up to 200 times or until convergence of the contour was achieved.

Based on the final masks, the shape complexity of each lesion was quantified using the perimeter-to-area ratio (*PAR*). This feature was defined as follows:

$$PAR = \frac{P}{A} \quad (2)$$

where *P* represents the length of the lesion boundary (perimeter) and *A* the enclosed lesion area. The perimeter was obtained from the binary mask contour length, while the area was calculated as the number of pixels within the mask, converted to mm<sup>2</sup> according to the pixel spacing in both imaging axes.

Higher *PAR* values reflect increased boundary irregularity, which may indicate invasive growth. This parameter is particularly useful for differentiating thyroid cancer subtypes, especially in identifying lesions with irregular, spiculated borders typical of PTC. In contrast, FTC and MTC more frequently exhibit smooth, well-defined borders [42–44].

### 2.3.2. Echogenicity and Internal Echotexture Features

- Echogenicity

To quantitatively assess the echogenicity of each focal thyroid lesion, RGB ultrasound images were converted to single-channel grayscale format and normalized to a [0, 1] intensity range. Using the corresponding binary segmentation mask, only the pixels within the lesion area were extracted, excluding the background and adjacent anatomical structures.

To minimize the influence of artifacts and technical inhomogeneities, pixels with extreme brightness values—above 0.85 (which may correspond to calcifications) and below 0.2 (potentially representing cystic or fluid-filled regions)—were excluded from further analysis.

For each lesion, the mean ( $\mu_{lesion}$ ) and median ( $M_{lesion}$ ) grayscale intensities were computed within the lesion mask. To account for inter-patient variability and reduce dependence on ultrasound system settings, these values were normalized to the mean intensity of the surrounding normal thyroid parenchyma ( $\mu_{thyroid}$ ):

$$NE_{mean} = \frac{\mu_{lesion}}{\mu_{thyroid}}, \quad NE_{median} = \frac{M_{lesion}}{\mu_{thyroid}} \quad (3)$$

where  $NE_{mean}$  and  $NE_{median}$  denote the normalized mean and median echogenicity, respectively.

Normalization provided a patient-specific reference, ensuring that relative rather than absolute echogenicity values were compared across cases. This step minimized bias from gain settings, probe type, or depth-related attenuation, making the echogenicity features more robust and reproducible across the study cohort.

- Internal Echotexture Features

To quantitatively assess the internal echotexture of each lesion, both global and local intensity variability were analyzed.

Global heterogeneity was estimated by computing the standard deviation of grayscale intensity values within the binary lesion mask:

$$\sigma = \sqrt{\frac{1}{N} \sum_{i=1}^N (I_i - \mu)^2} \quad (4)$$

where  $I_i$  denotes the intensity of the  $i$ -th pixel inside the lesion mask,  $\mu$  is the mean intensity of the lesion, and  $N$  is the total number of pixels in the lesion.

Local disorder was assessed using entropy. For each pixel, entropy was computed in a circular neighborhood of radius 0.7 mm:

$$E = - \sum_{k=1}^L p_k \log(p_k) \quad (5)$$

where  $p_k$  denotes the probability of intensity level  $k$  within the local neighborhood and  $L$  is the number of quantized intensity levels. The mean entropy value across all pixels within the lesion was then used as a marker of microstructural disorganization.

In addition to these intensity-based descriptors, second-order texture features were extracted from the gray-level co-occurrence matrix (GLCM). Given a displacement vector  $(\Delta x, \Delta y)$ , the normalized GLCM  $P(i, j)$  encodes the probability of finding a pair of gray levels,  $i$  and  $j$ , at that spatial offset. From this matrix, four features were computed:

$$Contrast = \sum_{i=1}^L \sum_{j=1}^L (i - j)^2 P(i, j) \quad (6)$$

measuring local brightness variations and sharp transitions,

$$Correlation = \sum_{i=1}^L \sum_{j=1}^L \frac{(i - \mu_i)(j - \mu_j) P(i, j)}{\sigma_i \sigma_j} \quad (7)$$

capturing linear dependencies between neighboring intensities, where  $\mu_i$ ,  $\mu_j$ ,  $\sigma_i$ , and  $\sigma_j$  are marginal means and standard deviations of the distribution,

$$Homogeneity = \sum_{i=1}^L \sum_{j=1}^L \frac{P(i, j)}{1 + |i - j|} \quad (8)$$

reflecting local uniformity, with higher values for smoother textures,

$$Energy = \sum_{i=1}^L \sum_{j=1}^L P(i, j)^2 \quad (9)$$

quantifying the degree of repetition in brightness patterns.

Higher values of standard deviation and entropy indicate increased echotextural inhomogeneity [45], whereas the GLCM-derived features capture more subtle patterns of pixel interrelationships. These quantitative descriptors may reveal microstructural differences among thyroid cancer subtypes, potentially supporting improved diagnostic discrimination [46–49].

### 2.3.3. Boundary Characteristics

- Assessment of Lesion Boundary Sharpness

To quantitatively evaluate the sharpness of lesion boundaries on B-mode ultrasound images, two complementary approaches were applied: (1) analysis of intensity gradients along the lesion contour and (2) evaluation of local intensity transitions across the lesion margin.

All grayscale images were first normalized relative to the mean echogenicity of adjacent normal thyroid parenchyma. For each point along the lesion contour, a margin region with a total width of 1 mm ( $\pm 0.5$  mm on each side) was defined, incorporating pixels both inside and outside the lesion. Descriptive statistics of these intensity changes—means, medians, and standard deviations—were computed to characterize the overall contrast along the boundary.

In addition, for each contour point, the normal direction,  $n$ , was estimated as the local gradient of the binary mask. Along this direction, a one-dimensional intensity profile,  $I(s)$ , was extracted, where  $s$  denotes the position in millimeters relative to the boundary ( $s = 0$  on the contour, negative values inside the lesion, positive values outside).

The local contrast at a given boundary point was then defined as follows:

$$C_{local} = \max_{s \in [-0.5, 0.5]} I(s) - \min_{s \in [-0.5, 0.5]} I(s) \quad (10)$$

For each lesion, the distribution of local contrast values across all contour points was summarized by its mean and standard deviation:

$$C = \frac{1}{M} \sum_{k=1}^M C_{local,k} \quad \sigma_C = \sqrt{\frac{1}{M} \sum_{k=1}^M (C_{local,k} - C)^2} \quad (11)$$

where  $M$  is the number of sampled contour points.

Together, these parameters—contrast statistics within the margin band and local contrast along normal profiles—provide a robust, device-independent assessment of boundary sharpness. Increased sharpness typically indicates well-delineated lesions (more common in FTC and MTC), whereas blurred or heterogeneous boundaries suggest infiltrative growth, which is often associated with PTC [50,51].

- Boundary Blurring—Kullback–Leibler Divergence Relative to Normal Parenchyma

To quantitatively assess the distinctiveness of lesion boundaries, the Kullback–Leibler (KL) divergence was computed between the local intensity distribution at the lesion margin and the intensity distribution within normal thyroid parenchyma. This parameter was used as a measure of lesion-to-background separability—higher values indicate a well-defined boundary with clearly distinguishable signal properties, whereas lower values suggest boundary blurring and similarity to the surrounding tissue.

For each lesion, a margin of 1 mm thickness was generated around the lesion contour, as defined by the segmentation mask. Grayscale images were previously normalized relative to the mean echogenicity of normal thyroid tissue to ensure comparability of intensity distributions across cases. The intensity distribution was used to estimate a probability distribution for two regions: (1) the margin zone surrounding the lesion boundary and (2) a reference region within the adjacent healthy parenchyma.

The KL divergence ( $D_{KL}$ ) was computed according to the following formula:

$$D_{KL}(P \parallel Q) = \sum_{i=1}^N P(i) \cdot \log \left( \frac{P(i)}{Q(i)} \right) \quad (12)$$

where  $P(i)$  denotes the probability of intensity in the  $i$ -th histogram bin within the margin region and  $Q(i)$  is the corresponding probability in the reference region.

Lower values of  $D_{KL}$  were interpreted as indicative of poor boundary distinctiveness, with intensity distributions closely resembling those of the surrounding parenchyma. In contrast, higher values reflected greater divergence between lesion and background tissue, suggesting a sharp, well-delineated border. This measure served as an objective and observer-independent parameter for characterizing the degree of lesion separability and for supporting the differentiation of focal thyroid lesions.

#### 2.3.4. Structural Features

- Micro- and Macrocalcifications

To quantitatively assess the presence of calcifications in focal thyroid lesions, a detailed analysis of binary masks corresponding to calcified areas and total lesion area was performed. Potential calcifications were identified on ultrasound images by applying an intensity threshold: regions with normalized grayscale values exceeding 0.85 were considered highly echogenic and classified as potential calcifications. Subsequently, morphological filtering was used to eliminate small artifacts that did not meet predefined morphological criteria. The resulting binary masks were automatically generated and then verified by visual comparison with B-mode images to confirm the correct localization of hyper-echoic foci.

For each lesion, the total lesion area ( $A_{lesion}$ ) and the area occupied by calcifications ( $A_{calc}$ ) were calculated in square millimeters, based on the true spatial resolution of the image (pixel dimensions in the X and Y axes). In previous studies, the distinction between micro- and macrocalcifications was defined using both 1 mm and 2 mm thresholds. In this study, an intermediate value was adopted as a compromise between these criteria [52,53]. Calcifications were further classified as microcalcifications ( $\leq 1.5$  mm) or macrocalcifications ( $> 1.5$  mm) according to their maximum transverse dimensions.

Two quantitative indicators were then computed: calcification density (for micro- or macrocalcifications), defined as the number of respective foci ( $N_{micro}$  and  $N_{macro}$ ) per unit area of the lesion:

$$D_{micro} = \frac{N_{micro}}{A_{lesion}} \quad D_{macro} = \frac{N_{macro}}{A_{lesion}} \quad (13)$$

and calcification area ratio, defined as the percentage of the lesion area occupied by calcifications:

$$R_{calc} = \frac{A_{calc}}{A_{lesion}} \times 100\%, \quad (14)$$

In addition, the spatial distribution of calcifications within the lesion was assessed. Based on the segmentation mask, the central region of the lesion was defined via morphological erosion of the full lesion mask, while the peripheral zone was defined as the difference between the original and eroded masks. This approach enabled differentiation between calcifications located near the lesion margin and those situated deeper within the lesion core. Such an analysis provided a more comprehensive, size-independent assessment of calcification patterns across focal thyroid lesions [52–54].

- Anechoic Areas (Cystic Components and Necrosis)

On B-mode ultrasound images, some focal thyroid lesions may contain anechoic regions, which appear as markedly hypoechoic areas within the tumor. The presence of such regions may indicate cystic components, liquefied necrosis, or fluid-filled zones within otherwise solid structures [55–57]. Therefore, one of the analyzed features was the presence and spatial characterization of anechoic areas within the lesion.

All ultrasound images were converted to grayscale and normalized relative to the mean intensity of normal thyroid parenchyma. Anechoic areas were defined as contiguous regions with intensity values below 0.2, located entirely within the lesion boundaries. To reduce the impact of noise and exclude isolated low-intensity pixels, morphological filtering was applied to retain only spatially coherent structures.

For each lesion, the total lesion area ( $A_{lesion}$ ) and the area of anechoic regions ( $A_{anechoic}$ ) were computed as described previously. The presence of anechoic areas was then quantified using the anechoic area ratio, defined as the percentage of the lesion area occupied by these regions:

$$R_{anechoic} = \frac{A_{anechoic}}{A_{lesion}} \times 100\% \quad (15)$$

#### 2.4. Statistical Analysis

All statistical analyses were performed using MATLAB R2023b (MathWorks, Natick, MA, USA). Prior to group comparisons, the distribution of continuous variables was assessed using the Shapiro–Wilk test, which is recommended as a sensitive method for evaluating normality in biomedical data [58]. Since most of the analyzed features did not meet the assumptions of normal distribution, the non-parametric Kruskal–Wallis test was used to compare the three histological subtypes: papillary (PTC), follicular (FTC), and medullary (MTC) thyroid carcinoma.

When significant differences were detected using the Kruskal–Wallis test, post hoc pairwise comparisons were performed using Dunn’s test, with significance level adjustment based on Šidák correction to control the risk of type I error associated with multiple comparisons [59,60].

A  $p$ -value  $< 0.05$  was considered statistically significant. The analysis included all quantitative ultrasound-based features, including geometric parameters, echogenicity, texture metrics, boundary characteristics, presence of calcifications, and anechoic components. Statistical results are reported as  $p$ -values for features showing significant intergroup differences.

#### 2.5. Multiparametric Classification Based on Quantitative Imaging Features

A three-class classification model based on the Random Forest algorithm (100 trees) was developed to differentiate between PTC, FTC, and MTC based on ultrasound-derived features. The dataset (244 images representing 122 lesions) was randomly divided into a training set (70%,  $n = 171$ ) and a test set (30%,  $n = 73$ ), with stratification to preserve class proportions. Each tree in the ensemble was trained on a bootstrap sample of the training set. At each node, a random subset of predictors (proportional to the square root of the total number of features) was considered for splitting. Splits were selected using the Gini impurity index, ensuring optimal separation of classes. Individual trees generated independent class predictions, and the final classification was determined by majority voting across all trees.

Model performance was evaluated using two strategies: (1) out-of-bag (OOB) validation, based on predictions from trees that did not include a given sample during training, and (2) an independent test set, providing an unbiased estimate of generalization performance [61–64].

After training the full model, feature importance was estimated using the OOB permutation method (OOBPermutedPredictorDeltaError). A feature was considered important if random permutation of its values increased classification error. Based on this ranking, the 10 most informative predictors were selected to construct a reduced model. The same training and validation procedures were applied to evaluate whether limiting the model to the most relevant features could preserve diagnostic performance while improving interpretability [65–68].

## 2.6. Software and Data Availability

All procedures for image analysis and feature extraction were implemented in MATLAB R2023b (MathWorks, Natick, MA, USA) using custom scripts. The processing pipeline included grayscale conversion, intensity normalization, segmentation refinement, quantitative feature computation, statistical analysis, and construction of a multiparametric classification model.

Anonymized B-mode ultrasound images and the corresponding segmentation masks used for quantitative analysis are available from the corresponding author upon reasonable request.

## 3. Results

### 3.1. Quantitative Evaluation of Single Ultrasound Features

#### 3.1.1. Morphological Feature Assessment: Shape and Complexity

The aspect ratio, describing the spatial orientation of the lesion, did not significantly differentiate thyroid cancer subtypes ( $p = 0.297$ ). The lowest median values were observed in MTC (0.78), while the highest were noted in FTC (0.98), which may reflect the more regular, oval shape of follicular tumors. In PTC, the widest range of values was noted, including cases with an aspect ratio  $> 1$ , potentially indicating a more vertical, infiltrative growth pattern.

The perimeter-to-area ratio, reflecting boundary complexity, significantly differentiated tumor subtypes (Kruskal–Wallis test:  $p < 0.0001$ ). Post hoc Dunn–Šidák analysis revealed significant differences between PTC vs. FTC ( $p < 0.0001$ ) and MTC vs. FTC ( $p = 0.0002$ ). The highest values were observed in the PTC group (median = 0.385), and the lowest in FTC (0.191), suggesting greater boundary irregularity in papillary carcinomas.

#### 3.1.2. Echogenicity and Intratumoral Texture Characteristics

Echogenicity showed significant differences among the three histological subtypes for both the mean ( $p = 0.0003$ ) and median values ( $p = 0.0002$ ). Post hoc analysis confirmed statistically significant differences for all pairwise comparisons: MTC vs. FTC ( $p \leq 0.0002$ ), MTC vs. PTC ( $p \approx 0.014$ – $0.016$ ), and PTC vs. FTC ( $p \approx 0.023$ – $0.035$ ). The lowest values were observed in MTC, and the highest in FTC. In contrast, the standard deviation of echogenicity did not significantly differ between groups ( $p = 0.1121$ ).

Local entropy, representing the degree of signal disorder within the lesion, also showed significance ( $p = 0.0360$ ), with a significant post hoc difference identified between MTC and FTC ( $p = 0.0486$ ). Other texture features based on the gray-level co-occurrence matrix (GLCM) did not reach statistical significance, although contrast approached the significance threshold ( $p = 0.0565$ ).

#### 3.1.3. Assessment of Tumor Margins

Intensity gradients and local contrast along the lesion boundary significantly differentiated between thyroid cancer subtypes. For the mean gradient, a significant difference was found between PTC and FTC ( $p = 0.0339$ ), while the gradient standard deviation differed significantly between PTC and FTC ( $p = 0.0014$ ) and between MTC and FTC ( $p = 0.0163$ ). Similarly, mean profile intensity differed between PTC and FTC ( $p = 0.0430$ ), and the standard deviation of profile intensities showed differences between PTC and FTC ( $p = 0.0009$ ) as well as between MTC and FTC ( $p = 0.0180$ ). The lack of significant differences between PTC and MTC for these features suggests that boundary-related parameters are most effective in distinguishing FTC from the other subtypes.

The Kullback–Leibler divergence, which quantifies the distinction between the lesion and the surrounding thyroid parenchyma, also showed significant differences ( $p = 0.0049$ ), with post hoc analysis revealing a significant difference between MTC and PTC ( $p = 0.0165$ ).

#### 3.1.4. Internal Composition and Calcification Patterns

Macrocalcifications and the percentage of the tumor area occupied by calcifications significantly differed across thyroid cancer subtypes ( $p = 0.0112$  and  $p = 0.0435$ , respectively). Post hoc Dunn–Šidák analysis revealed significant differences between PTC and FTC ( $p \approx 0.0081$ – $0.0399$ ). In contrast, microcalcifications did not significantly differ between groups ( $p = 0.7264$ ).

Peripheral calcifications showed highly significant differences between subtypes ( $p < 0.00001$ ). The highest median count was observed in FTC (17), markedly exceeding the values for PTC and MTC (median = 3). These differences were statistically significant for the FTC–PTC and FTC–MTC comparisons ( $p < 0.0001$ ).

Anechoic (cystic) areas were most frequently observed in FTC lesions; however, the Kruskal–Wallis test did not show significant differences between groups ( $p = 0.7902$ ), limiting the diagnostic utility of this feature as a standalone parameter.

#### 3.2. Comparative Evaluation of Individual Quantitative Ultrasound Features

All analyzed ultrasound features are summarized in Table 2, which presents the statistical significance of differences between the three thyroid carcinoma subtypes (PTC, FTC, and MTC) for each pairwise comparison. Statistically significant differences are marked with an “X”. Based on this comparative analysis of individual imaging features, it can be observed that only certain parameters showed significant differences between selected group pairs. For instance, the mean and median echogenicity, as well as features describing boundary complexity, significantly differentiated PTC from FTC, as well as MTC from FTC. In contrast, other features—such as the presence of anechoic areas or selected GLCM-based texture metrics—did not independently show significant differences between the subtypes.

To reduce the impact of multicollinearity among features, a correlation analysis was performed for all extracted imaging parameters. Pearson correlation coefficients were calculated for each pair of features. Several pairs of variables exhibited strong linear correlations ( $|r| > 0.9$ ), which may have led to redundancy and disproportionately influenced the classification model.

To minimize this redundancy, only one parameter was retained from each pair of highly correlated features. When selecting which variable to keep, priority was given to those that demonstrated statistical significance in the three-group analysis and offered greater clinical interpretability. Special emphasis was placed on features that could be more readily understood and applied in clinical practice.

This approach allowed the number of input variables to be reduced to 14 independent features, contributing to lower model variance and improved interpretability. The final set of features used for model construction is presented in Table 3.

**Table 2.** Results of a two-step statistical analysis. First, the Kruskal–Wallis test was applied to assess overall differences in the distribution of quantitative ultrasound features among thyroid cancer subtypes. Reported  $p$ -values in this column indicate the global significance level for each parameter. For features with statistically significant results ( $p < 0.05$ ), post hoc Dunn–Šidák pairwise comparisons were performed to identify specific group differences, while for non-significant parameters post hoc comparisons were not applicable (marked as “–”). Statistically significant  $p$ -values are shown in bold.

Parameter Group	Parameter	Kruskal–Wallis ( $p$ )	Post Hoc Dunn–Šidák Comparisons ( $p$ )		
			PTC vs. FTC	PTC vs. MTC	MTC vs. FTC
Morphological Features	Aspect ratio	0.297	–	–	–
	Perimeter-to-area ratio	<b>&lt;0.0001</b>	<b>&lt;0.0001</b>	0.0674	<b>0.0002</b>
Internal Architecture	Echogenicity (mean)	<b>0.0003</b>	<b>0.0234</b>	<b>0.0141</b>	<b>0.0002</b>
	Echogenicity (median)	<b>0.0002</b>	<b>0.0352</b>	<b>0.0162</b>	<b>0.0019</b>
	Echogenicity (std)	0.1121	–	–	–
	Local entropy (mean)	<b>0.0360</b>	0.2147	0.1863	<b>0.0486</b>
	Local entropy (std)	0.0673	–	–	–
	Contrast (mean)	0.0565	–	–	–
	Correlation (mean)	0.1569	–	–	–
	Homogeneity (mean)	0.9859	–	–	–
	Energy (mean)	0.7586	–	–	–
Margin Assessment	Gradient (mean)	<b>0.0400</b>	<b>0.0339</b>	0.9464	0.2416
	Gradient (std)	<b>0.0021</b>	<b>0.0014</b>	0.8999	<b>0.0163</b>
	Profile (mean)	<b>0.0014</b>	<b>0.0430</b>	0.8193	0.3867
	Profile (std)	0.0443	<b>0.0009</b>	0.9909	<b>0.0180</b>
	KL divergence	<b>0.0049</b>	0.1268	<b>0.0165</b>	0.6894
Structural Features	Microcalcification density	0.7264	–	–	–
	Macrocalcification density	<b>0.0112</b>	<b>0.0081</b>	0.9811	0.0834
	Calcified area %	<b>0.0435</b>	<b>0.0399</b>	0.9989	0.1074
	Peripheral calcification	<b>&lt;0.0001</b>	<b>&lt;0.0001</b>	<b>0.0127</b>	<b>&lt;0.0001</b>
	Cystic area %	0.7902	–	–	–

**Table 3.** Final set of imaging features after collinearity reduction—results of the Kruskal–Wallis test with Dunn–Šidák post hoc comparisons for parameters differentiating thyroid cancer subtypes (PTC, FTC, and MTC). Exact  $p$ -values are provided; statistically significant differences ( $p < 0.05$ ) are shown in bold.

Parameter Group	Parameter	Kruskal–Wallis ( $p$ )	Post Hoc Dunn–Šidák Comparisons ( $p$ )		
			PTC vs. FTC	PTC vs. MTC	MTC vs. FTC
Morphological Features	Aspect ratio	0.297	–	–	–
	Perimeter-to-area ratio	<b>&lt;0.0001</b>	<b>&lt;0.0001</b>	0.0674	<b>0.0002</b>
Internal Architecture	Echogenicity (mean)	<b>0.0003</b>	<b>0.0234</b>	<b>0.0141</b>	<b>0.0002</b>
	Echogenicity (std)	0.1121	–	–	–
	Local entropy (mean)	<b>0.0360</b>	0.2147	0.1863	<b>0.0486</b>
	Contrast (mean)	0.0565	–	–	–
Margin Assessment	Gradient (std)	<b>0.0021</b>	<b>0.0014</b>	0.8999	<b>0.0163</b>
	Profile (mean)	<b>0.0014</b>	<b>0.0430</b>	0.8193	0.3867
	KL divergence	<b>0.0049</b>	0.1268	<b>0.0165</b>	0.6894
Structural Features	Microcalcification density	0.7264	–	–	–
	Macrocalcification density	<b>0.0112</b>	<b>0.0081</b>	0.9811	0.0834
	Calcified area %	<b>0.0435</b>	<b>0.0399</b>	0.9989	0.1074
	Peripheral calcification	<b>&lt;0.0001</b>	<b>&lt;0.0001</b>	<b>0.0127</b>	<b>&lt;0.0001</b>
	Cystic area %	0.7902	–	–	–

### 3.3. Classification Model Based on Full Feature Set

The classification model based on the full set of 14 imaging features (Table 3) achieved an overall classification accuracy of 89.3%. However, class-wise performance varied in predicting specific tumor subtypes.

Class-wise performance, recalculated to match class sizes is summarized in Table 4 and revealed the following: for FTC, a precision of 75.0%, a recall of 64.3%, and an F1-score of 69.1%; for MTC, a precision of 88.2%, a recall of 83.3%, and an F1-score of 85.7%; and for PTC, a precision of 92.4%, a recall of 94.4%, and an F1-score of 93.4%.

**Table 4.** Classification performance metrics for the model built using the full feature set.

Class	Precision (%)	Recall (%)	F1-Score (%)
FTC	75.0	64.3	69.1
MTC	88.2	83.3	85.7
PTC	92.4	94.4	93.4

The confusion matrix presented in Table 5 shows that the most common misclassifications involved FTC being labeled as PTC and MTC being labeled as PTC, which may be attributed to overlapping morphological characteristics.

**Table 5.** The confusion matrix for the model built using the full feature set.

True Class/ Predicted Class	FTC (pred)	MTC (pred)	PTC (pred)
FTC (true)	64.3	0.0	35.7
MTC (true)	0.0	83.3	16.7
PTC (true)	3.3	2.2	94.4

### 3.4. Feature Importance and Reduced Feature Model

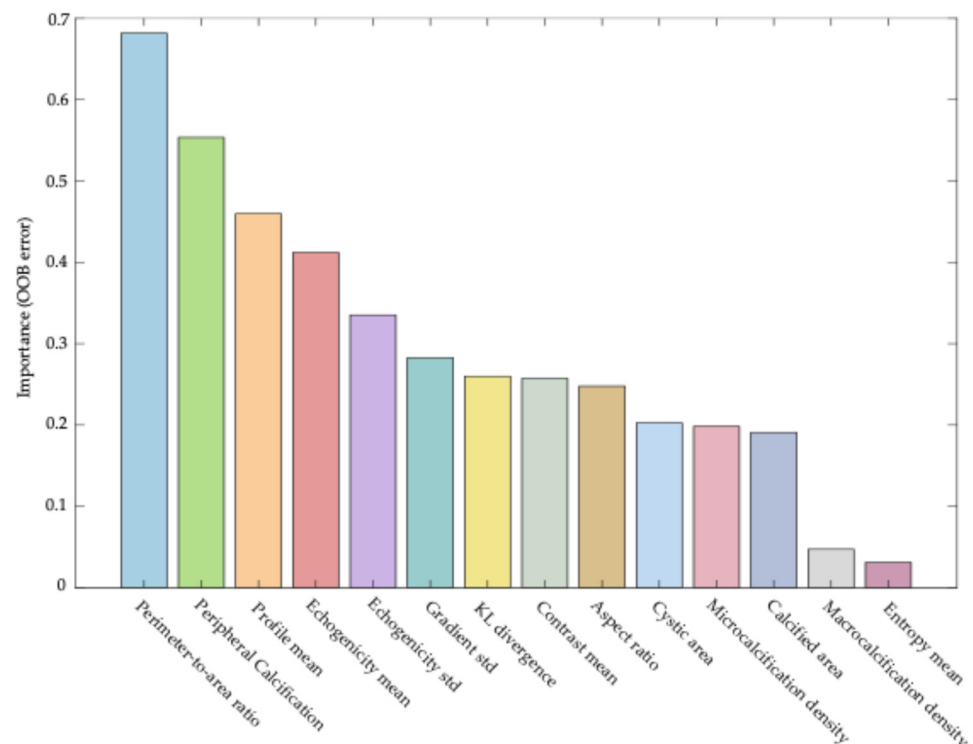
Following the evaluation of the classification performance of the model based on the full set of features, an analysis of predictor importance was conducted. The permutation-based importance metric (OOBPermutedPredictorDeltaError) was used, which measures the increase in out-of-bag (OOB) error after randomly permuting the values of a given feature.

This analysis enabled ranking of the features according to their impact on the model's predictive accuracy. High values of the permutation importance index indicated a significant contribution of the feature to classification decisions, whereas values close to zero or negative values suggested limited or no diagnostic relevance. The results of this analysis are presented as a bar plot in Figure 1, with features ordered in descending order of importance.

Among the features with the highest information value in the classification model were both morphological parameters, such as the perimeter-to-area ratio (feature #1), and features related to calcification, including the presence of peripheral calcifications (feature #2). Parameters associated with lesion echogenicity also played a significant role, including mean echogenicity (feature #4) and the mean boundary profile value (feature #3).

While some individual features demonstrated clear dominance in importance, the results highlight the value of a multidimensional approach, in which complementary information derived from different image aspects (morphology, texture, echogenicity, and calcifications) collectively contributes to accurate differentiation of thyroid cancer histological subtypes.

Based on these findings, the 10 most important predictors were selected to construct a simplified classification model. The aim of this step was to evaluate whether reducing the number of input variables could maintain high diagnostic accuracy while simplifying the model's structure and improving interpretability.



**Figure 1.** Importance of features in the Random Forest model. The vertical axis represents the increase in out-of-bag (OOB) classification error following random permutation of each feature, reflecting its relative contribution to model performance. Features are ranked in descending order of importance.

The model based on the top 10 features achieved a classification accuracy of 91.8%, indicating strong predictive performance even with a reduced feature set. For the FTC class, the model reached a precision of 66.7% and a recall of 85.7%, resulting in an F1-score of 70.0%. In the case of MTC, the precision and recall were 70.0% and 77.8%, respectively, yielding an F1-score of 73.6%. For PTC, the model demonstrated very high recall (95.6%) and precision (96.1%), with an F1-score of 94.0%. Detailed results are presented in Table 6.

**Table 6.** Classification performance metrics for the model built using the 10 most informative features.

Class	Precision (%)	Recall (%)	F1-Score (%)
FTC	66.7	85.7	75.0
MTC	70.0	77.8	73.6
PTC	96.6	95.6	96.1

The confusion matrix presented in Table 7 shows that the most common misclassifications involved FTC cases being predicted as PTC (14.3%) and MTC cases being predicted as PTC (22.2%). This may be attributed to partial overlap in imaging features between these thyroid cancer subtypes.

**Table 7.** Confusion matrix (%)—classification based on the 10 most important features.

True Class/ Predicted Class	FTC (pred)	MTC (pred)	PTC (pred)
FTC (true)	85.7	0.0	14.3
MTC (true)	0.0	77.8	22.2
PTC (true)	1.1	3.3	95.6

A comparison of confusion matrices between the full model (based on 14 features) and the reduced model (using the 10 most important predictors) indicates comparable classification performance for both approaches. For the FTC class, the reduced model even achieved slightly higher classification accuracy (85.7% vs. 77.8% in the full model), suggesting that reducing the number of input variables did not impair recognition of this class. A slight decrease was observed for MTC, with classification accuracy decreasing from 83.3% to 77.8%, although misclassifications as PTC increased slightly (22.2% vs. 16.7%). For the PTC class, the full model showed slightly better performance (94.4% vs. 95.6%), but this difference had a negligible impact on overall classification accuracy.

These findings confirm that the simplified classification model, despite using fewer features, maintained high predictive performance—outperforming the full model in some cases. The slight decrease in accuracy for PTC was offset by improved recognition of FTC and MTC cases. Therefore, reducing the number of input variables may not only enhance the model's interpretability but also increase its generalizability, particularly for underrepresented tumor subtypes.

#### 4. Discussion

The results of the conducted analyses confirm that a quantitative approach to ultrasound assessment of thyroid nodules, based on objective and standardized B-mode parameters, enables effective differentiation of the three main histological subtypes of thyroid carcinoma: papillary (PTC), follicular (FTC), and medullary (MTC). The Random Forest model constructed using the full set of 14 imaging features achieved a high classification accuracy of 89.3%, which was preserved in the simplified version based on only the 10 most important predictors (91.8%). This confirms that a well-selected, reduced set of features can provide equally high diagnostic performance while improving model interpretability and reducing the risk of overfitting.

In the permutation-based importance analysis, four predictors were identified as having the highest diagnostic relevance, each representing a distinct category of quantitative ultrasound features: perimeter-to-area ratio (morphology), peripheral calcification (structural features), profile mean (boundary sharpness), and echogenicity mean (internal echogenicity). Each of these features reflects a different aspect of the nodule's sonographic appearance, namely, irregular margins, the presence of peripheral calcifications, contrast along the margin, and internal texture, respectively. This complementary integration of multiple imaging domains contributed to the high classification performance and enhanced the interpretability of the model in relation to the underlying histopathological differences among thyroid cancer subtypes.

In this discussion, only selected features from the broader set of evaluated parameters are highlighted, as detailed elaboration on all of them would exceed the scope of this article.

Among the listed predictors, perimeter-to-area ratio demonstrated the highest predictive importance score (0.70), indicating its key role in the classification process. Its diagnostic value stems from its biological interpretability. Papillary thyroid carcinoma (PTC) typically exhibits infiltrative growth and papillary architecture, resulting in irregular, spiculated margins and elevated perimeter-to-area ratios [69–71]. In contrast, follicular thyroid carcinoma (FTC) more often presents as encapsulated lesions with smooth margins, leading to lower perimeter-to-area values [70,72,73]. Medullary thyroid carcinoma (MTC) shows a more variable morphology but often also smooth margins [70,74].

Regardless of these biological differences, the effectiveness of this approach relies not only on the choice of the metric itself but also on the methodology used to extract it. The accuracy of margin analysis is particularly sensitive to the method of segmentation, especially in the case of nodules with ill-defined or infiltrative borders. Unlike fully automated

approaches, which frequently struggle with precise segmentation in low-contrast areas, semi-automated and iterative techniques allow for more reliable and reproducible results. Recent studies have demonstrated that integrating methods such as active contours, morphological filtering, and specialized boundary-sensitive modules significantly improves segmentation performance, especially in challenging cases [75–79]. The combination of boundary-based features with morphological operations such as dilation and erosion enables more accurate contour delineation, even in the presence of blurred edges [75,77,79].

Importantly, unlike subjective features such as “irregular margins” assessed by radiologists, the perimeter-to-area ratio is an entirely objective and reproducible metric. This makes it particularly attractive in the context of malignancy risk stratification and automation. It may assist clinicians in biopsy decision-making and serve as a core component of future AI-based decision support systems.

Due to its biological relevance and interpretability, the perimeter-to-area ratio can also serve as a conceptual bridge between traditional image-based diagnostics and modern machine learning algorithms. Incorporating this feature into predictive models enhances both diagnostic accuracy and the transparency of the decision-making process, which is essential for clinical acceptance of AI-assisted diagnostic tools.

Peripheral or rim calcifications have a distinct diagnostic value, as their quantitative assessment, such as the number of high-intensity pixels along the lesion contour, reflects the presence of calcifications at the tumor border. In the TI-RADS and EU-TIRADS classifications, it is emphasized that peripheral calcifications, particularly those with an “eggshell” appearance and interrupted continuity, may be associated with an increased risk of malignancy, whereas regular, continuous rim calcifications are typically found in benign lesions [9–11,80–82]. Studies have shown that central microcalcifications are most commonly observed in papillary thyroid carcinoma, while peripheral macrocalcifications may occur in papillary, follicular, and medullary thyroid carcinomas, with their prognostic significance depending on morphology and continuity [80,82–84]. Interrupted rim calcifications are associated with a higher risk of malignancy and may indicate an infiltrative growth pattern, whereas continuous “eggshell”-type calcifications are characteristic of benign lesions [80–82].

Statistical analyses confirm that the presence of peripheral calcifications—particularly when combined with other suspicious features such as solid composition or irregular margins—is associated with an increased risk of malignancy [81,82]. However, the presence of macrocalcifications or peripheral calcifications alone, in the absence of other suspicious characteristics, does not necessarily indicate a high risk of cancer [85–87]. Recent studies suggest that automated, quantitative analysis of calcifications—including their number, distribution, intensity, and distance from the lesion margin—may aid in differentiating benign from malignant nodules and in predicting metastatic risk, especially in papillary thyroid carcinoma [81,85,88]. Deep learning-based models allow for automatic detection and assessment of calcifications, and the parameters obtained show strong agreement with expert radiologist evaluation and significant prognostic value, although their use in routine clinical practice remains limited. The integration of such tools with AI-based diagnostic systems has the potential to substantially improve the classification of lesions with ambiguous echogenicity or architecture, thereby supporting clinical decision-making [85,88].

The parameter profile mean, which describes the average contrast at the boundary of a lesion relative to the surrounding thyroid parenchyma, plays a key role in differentiating nodules based on the architecture of their margins. A high value of this metric reflects a sharp, well-defined transition between the lesion and adjacent tissue, corresponding to the “sharp margins” criterion in the TI-RADS and EU-TIRADS classifications, where indistinct or irregular borders are associated with a higher risk of malignancy [9–11]. Nu-

merous studies have confirmed that irregular, poorly defined margins are characteristic of infiltrative lesions such as papillary or medullary thyroid carcinoma, whereas benign lesions and certain subtypes of follicular carcinoma typically exhibit smooth, encapsulated contours [89–91]. Intensity profile analysis at the interface between parenchyma and lesion enables a quantitative assessment of these features, which may provide valuable insights into the growth pattern and malignant potential of the nodule [90,92,93]. Automated computation of the profile mean based on averaged intensity profiles around the lesion contour ensures high reproducibility and eliminates subjective variability, thereby supporting standardization and integration with artificial intelligence-based diagnostic tools [92,94].

In the presented study, the lowest values of echogenicity were observed in medullary thyroid carcinoma (MTC), which aligns with findings from previous reports [13,95]. For follicular thyroid carcinoma (FTC), the literature suggests a generally higher echogenicity, although clear comparative data remain limited [96,97]. Hypoechogenicity of a thyroid nodule is a well-established indicator of malignancy risk, particularly in papillary thyroid carcinoma (PTC), as demonstrated by both ultrasound and histopathological studies [13,96,97]. Ultrasonographic studies have shown that PTC is most commonly hypoechoic, while MTC may exhibit either low or mixed echogenicity. FTC, along with other follicular-patterned lesions, more frequently demonstrates echogenicity similar to that of the surrounding thyroid or only mildly reduced [13,95–97].

In the presented classification model, the echogenicity mean was computed using a fully automated method that included local normalization relative to the background tissue. This approach enhanced the robustness and reproducibility of the measurement, reducing the influence of inter-device variability [98].

The results further support the effectiveness of multiparametric models in differentiating thyroid focal lesions. Consistent with the literature, combining morphological, textural, and intensity-based features improves both sensitivity and specificity, while also increasing model resilience to noise and variability across different ultrasound systems [99–103]. Previous studies have reported high predictive performance for such hybrid models—for instance, Random Forest classifiers achieving accuracies of up to 96.1% [100] and other advanced algorithms exceeding 95% [100,103,104]. An important advantage of such models is the low correlation among key predictors, which limits redundancy and enhances classification efficiency [100,102]. Moreover, these models have demonstrated strong performance both with the full feature set and after dimensionality reduction, confirming their practical applicability in clinical diagnostics.

Recent advances in deep learning have introduced novel architectures for image analysis, including attention-based multiview frameworks [105], style-contrastive networks for content–style disentanglement [106], and transformer-based weakly supervised segmentation with adversarial training [107]. While these methods achieve state-of-the-art performance in various computer vision domains, their application in endocrine oncology remains limited, primarily due to the lack of clinical interpretability and the challenges in validating such black-box models. In contrast, the quantitative ultrasound features employed in our study are transparent, biologically interpretable, and directly linked to histopathological correlates, thereby supporting reproducibility and clinical trust.

Unlike deep learning methods, which are often perceived as opaque “black-box” systems, the approach presented here offers complete transparency in both analytical processing and interpretation of results. Each feature was designed based on a biologically grounded rationale, allowing for clinical interpretation of the measured values and supporting transparency in the decision-making process.

It is also worth noting that much of the existing literature has focused on binary classification tasks—distinguishing benign from malignant lesions—without addressing

finer distinctions among thyroid cancer subtypes [108–110]. In contrast, the present study tackled a three-class classification problem involving PTC, FTC, and MTC subtypes that differ not only in prognosis but also in architectural patterns, morphology, and biological behavior. Therefore, the proposed model aligns with current trends in quantitative ultrasound analysis while extending their application to more complex and clinically meaningful classification tasks.

Despite the high classification performance and the use of a fully quantitative and transparent approach, this study has several limitations. First, the analysis was conducted retrospectively and was based on data from a single institution, which may limit the generalizability of the findings across more diverse patient populations. Second, although the segmentation method was optimized and semi-automated, it still required expert supervision, which could affect reproducibility in settings with varying levels of operator experience. Third, the classification model was developed using B-mode ultrasound images acquired in two orthogonal planes (transverse and longitudinal). While this enabled the extraction of key morphological and structural features, three-dimensional ultrasound data were not included, which might have provided a more comprehensive representation of tumor architecture—particularly with respect to irregular margins, spatial distribution of calcifications, and internal heterogeneity.

In addition, the robustness of the proposed method to variations in ultrasound equipment, acquisition protocols, and device manufacturers has not yet been validated. Although Random Forest classifiers are relatively computationally efficient, the complete pipeline—including preprocessing and semi-automated segmentation—may impose additional computational demands, which could limit practical deployment in resource-constrained healthcare environments.

Future research should therefore address these challenges by including larger and more diverse multi-center cohorts, validating the approach across different ultrasound platforms, incorporating three-dimensional data, and optimizing computational efficiency. Such efforts will be crucial to ensure the broad applicability and real-world feasibility of quantitative ultrasound as a clinically trusted decision-support tool.

## 5. Conclusions

The application of quantitative ultrasound, supported by computational methods and machine learning, enables effective differentiation of the three major histological subtypes of thyroid cancer: papillary (PTC), follicular (FTC), and medullary (MTC). The Random Forest model, built upon selected morphological, structural, and textural features, achieved high classification accuracy (91.5%) while maintaining transparency and clinical interpretability.

The high sensitivity observed for PTC reflects the model's strong performance in detecting the most common thyroid cancer subtype, which is of substantial clinical relevance. At the same time, the high precision in classifying FTC and MTC suggests a low rate of misclassification for these less common but potentially more aggressive subtypes. The limited overlap in misclassifications between FTC and MTC, together with the model's robust diagnostic performance, highlights the potential of this approach as a clinically useful decision-support tool.

Validation of these findings requires further studies involving larger, more diverse cohorts and multicenter external validation to assess model performance in real-world settings. Integrating semi-automated segmentation and classification methods into the diagnostic workflow for thyroid nodules could reduce the number of unnecessary biopsies, support the implementation of personalized medicine strategies, and improve overall patient care efficiency.

**Author Contributions:** Conceptualization, H.P.W.; methodology, H.P.W.; software, H.P.W. and P.K.; validation, H.P.W.; formal analysis, H.P.W.; investigation, K.D.S. and A.Ż.; resources, K.D.S. and A.Ż.; data curation, K.D.S., A.Ż. and H.P.W.; writing—original draft preparation, H.P.W.; writing—review and editing, H.P.W., K.D.S., A.Ż., J.L. and P.K.; visualization, H.P.W.; supervision, H.P.W.; project administration, J.L.; funding acquisition, J.L.; clinical investigation and patient recruitment, M.D., K.D.S. and A.Ż. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research and publication was co-financed from the state budget under the program of the Minister of Education and Science, Poland, called “Science for Society II”, project number NdS-II/SP/0189/2024/01; amount of co-financing: PLN 1,000,000, total project value: PLN 1,000,000.

**Institutional Review Board Statement:** The study was conducted in accordance with the Declaration of Helsinki and approved by the Bioethics Committee of the Maria Skłodowska-Curie National Research Institute of Oncology in Warsaw (83/2021 and 30 November 2021).

**Informed Consent Statement:** Informed consent was obtained from all subjects involved in the study.

**Data Availability Statement:** Anonymized B-mode ultrasound images and the corresponding segmentation masks used for quantitative analysis are available from the corresponding author upon reasonable request.

**Conflicts of Interest:** The authors declare no conflicts of interest.

## References

1. Sung, H.; Ferlay, J.; Siegel, R.L.; Laversanne, M.; Soerjomataram, I.; Jemal, A.; Bray, F. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin.* **2021**, *71*, 209–249. [\[CrossRef\]](#)
2. Didkowska, J.; Wojciechowska, U.; Olasek, P.; Michałek, I.; Ciuba, A. *Nowotwory Złośliwe w Polsce w 2021 Roku*; Krajowy Rejestr Nowotworów, Narodowy Instytut Onkologii: Warszawa, Poland, 2023. (In Polish)
3. Cabanillas, M.E.; McFadden, D.G.; Durante, C. Thyroid cancer. *Lancet* **2016**, *388*, 2783–2795. [\[CrossRef\]](#)
4. Haugen, B.R.; Alexander, E.K.; Bible, K.C.; Doherty, G.M.; Mandel, S.J.; Nikiforov, Y.E.; Pacini, F.; Randolph, G.W.; Sawka, A.M.; Schlumberger, M.; et al. 2015 American Thyroid Association Management Guidelines for Adult Patients with Thyroid Nodules and Differentiated Thyroid Cancer. *Thyroid* **2016**, *26*, 1–133. [\[CrossRef\]](#)
5. Nikiforov, Y.E.; Biddinger, P.W.; Thompson, L.D.; Lester, D.R. *Diagnostic Pathology and Molecular Genetics of the Thyroid*, 2nd ed.; Wolters Kluwer: Philadelphia, PA, USA, 2012.
6. Wells, S.A., Jr.; Asa, S.L.; Dralle, H.; Elisei, R.; Evans, D.B.; Gagel, R.F.; Lee, N.; Machens, A.; Moley, J.F.; Pacini, F.; et al. Revised American Thyroid Association guidelines for the management of medullary thyroid carcinoma. *Thyroid* **2015**, *25*, 567–610. [\[CrossRef\]](#)
7. Roman, S.; Lin, R.; Sosa, J.A. Prognosis of medullary thyroid carcinoma: Demographic, clinical, and pathologic predictors of survival. *Arch. Surg.* **2005**, *140*, 891–897.
8. Pacini, F.; Castagna, M.G.; Brilli, L.; Pentheroudakis, G. Medullary thyroid cancer: ESMO Clinical Practice Guidelines. *Ann. Oncol.* **2010**, *21* (Suppl. S5), v219–v225. [\[CrossRef\]](#) [\[PubMed\]](#)
9. Tessler, F.N.; Middleton, W.D.; Grant, E.G.; Hoang, J.K.; Berland, L.L.; Teefey, S.A.; Cronan, J.J.; Beland, M.D.; Desser, T.S.; Frates, M.C.; et al. ACR Thyroid Imaging Reporting and Data System (TI-RADS): White paper of the ACR TI-RADS committee. *J. Am. Coll. Radiol.* **2017**, *14*, 587–595. [\[CrossRef\]](#) [\[PubMed\]](#)
10. Russ, G.; Bonnema, S.J.; Erdogan, M.F.; Durante, C.; Ngu, R.; Leenhardt, L. European Thyroid Association Guidelines for Ultrasound Malignancy Risk Stratification of Thyroid Nodules in Adults: The EU-TIRADS. *Eur. Thyroid J.* **2017**, *6*, 225–237. [\[CrossRef\]](#)
11. Migda, B.; Migda, M.; Migda, M.; Dobruch-Sobczak, K.; Słowińska-Klencka, D.; Woliński, K. EU-TIRADS-PL—Polskie wytyczne ultrasonograficznej stratyfikacji ryzyka złośliwości zmian ogniskowych tarczycy. *Endokrynol. Pol.* **2022**, *73*, 18–37. (In Polish)
12. Zhao, J.; Zheng, X.; Gao, M.; Zhang, S.; Yun, X.; Chi, J.; Xu, G. Ultrasound features of medullary thyroid carcinoma as predictors of biological behavior. *Cancer Imaging* **2021**, *21*, 33. [\[CrossRef\]](#) [\[PubMed\]](#)
13. Zhang, D.; Yang, F.; Wang, Y.; Mu, J.; Wei, X.; Wei, X. Ultrasonographic features of thyroid carcinoma of different sizes: Comparison between medullary thyroid carcinomas and papillary thyroid carcinomas. *Zhonghua Zhong Liu Za Zhi* **2024**, *46*, 133–139. [\[PubMed\]](#)
14. Lei, R.; Wang, Z.-D.; Qian, L. Ultrasonic Characteristics of Medullary Thyroid Carcinoma. *Ultrasound Q.* **2021**, *37*, 154–160. [\[CrossRef\]](#) [\[PubMed\]](#)
15. Hughes, N.M.; Nae, A.; Barry, J.; Fitzgerald, B.; Feeley, L.; Sheahan, P. Sonographic differences between conventional and follicular variant papillary thyroid carcinoma. *Eur. Arch. Otorhinolaryngol.* **2017**, *274*, 2021–2028. [\[CrossRef\]](#) [\[PubMed\]](#)

16. Hekimsoy, İ.; Ertan, Y.; Serin, G.; Karabulut, A.K.; Özbek, S. Comparison of ultrasound findings of papillary thyroid carcinoma subtypes based on the 2022 WHO classification of thyroid neoplasms. *Front. Endocrinol.* **2024**, *14*, 1434787. [\[CrossRef\]](#)
17. Chen, X.; Gao, X.; Ma, X.; Wei, B.; Bai, M. Efficacy of color Doppler ultrasound signs combined with serum tumor-specific growth factor in the diagnosis of differentiated thyroid cancer. *Am. J. Transl. Res.* **2024**, *16*, 3654–3666. [\[CrossRef\]](#)
18. Jiang, S.; Xie, Q.; Li, N.; Chen, H.; Chen, X. Modified models for predicting malignancy using ultrasound characters have high accuracy in thyroid nodules with small size. *Front. Mol. Biosci.* **2021**, *8*, 752417. [\[CrossRef\]](#)
19. Liu, M.; Liu, Z.; Hou, Y.; Men, Y.; Zhang, Y.; Gao, L.; Liu, H. Ultrasonographic characteristics of medullary thyroid carcinoma: A comparison with papillary thyroid carcinoma. *Oncotarget* **2017**, *8*, 27520–27528. [\[CrossRef\]](#)
20. Xu, Y.; Pi, J.; Jinghu, Y.; Wang, X.; Xu, D.; Liu, J. Diagnostic Efficiency of ACR-TIRADS Score for Differentiating Benign and Malignant Thyroid Nodules of Various Pathological Types. *Med. Sci. Monit.* **2024**, *30*, e943228. [\[CrossRef\]](#)
21. Solymosi, T.; Hegedüs, L.; Bodor, M.; Nagy, E. EU-TIRADS-Based omission of fine-needle aspiration and cytology from thyroid nodules overlooks a substantial number of follicular thyroid cancers. *Int. J. Endocrinol.* **2021**, *185*, 193–200. [\[CrossRef\]](#)
22. Ni, X.; Xu, S.Y.; Zhang, B.Y.; Zhan, W.; Zhou, W. Clinical and sonographic features of noninvasive follicular thyroid neoplasm with papillary-like nuclear features. *Ultrasound Q.* **2022**, *38*, 44–50. [\[CrossRef\]](#)
23. Dobruch-Sobczak, K.; Gumińska, A.; Bakula-Zalewska, E.; Mlosek, K.; Slapa, R.Z.; Wareluk, P.; Krauze, A.; Ziemiecka, A.; Migda, B.; Jakubowski, W.; et al. Shear wave elastography in medullary thyroid carcinoma diagnostics. *J. Ultrason.* **2015**, *15*, 358–367. [\[CrossRef\]](#)
24. Zhang, D.; Yang, F.; Hou, W.; Wang, Y.; Mu, J.; Wang, H.; Wei, X. Ultrasonic radiomics in predicting pathologic type for thyroid cancer: A preliminary study. *Front. Endocrinol.* **2025**, *16*, 1428888.
25. Kim, S.-H.; Kim, B.; Jung, S.; Lee, J.-W.; Yang, P.-S.; Kang, B.; Lim, H.; Kim, J.-Y.; Whang, I.; Kwon, H.; et al. Ultrasonographic findings of medullary thyroid carcinoma: A comparison with papillary thyroid carcinoma. *Korean J. Radiol.* **2009**, *10*, 101–105. [\[CrossRef\]](#) [\[PubMed\]](#)
26. Yang, G.; Fried, K.; Scognamiglio, T. Sonographic and cytologic differences of NIFTP from infiltrative or invasive encapsulated follicular variant papillary thyroid carcinoma. *Diagn. Cytopathol.* **2017**, *45*, 533–541. [\[CrossRef\]](#)
27. Ng, S.; Kuo, S.; Hua, C.; Huang, B.; Chiang, K.; Chu, Y.Y.; Hsueh, C.; Lin, J.D. Differentiation of the follicular variant of papillary thyroid carcinoma from classic papillary thyroid carcinoma: An ultrasound analysis and complement to fine-needle aspiration cytology. *J. Ultrasound Med.* **2018**, *37*, 667–674. [\[CrossRef\]](#)
28. Slabaugh, G.; Beltran, L.; Rizvi, H.; Deloukas, P.; Marouli, E. Applications of machine and deep learning to thyroid cytology and histopathology: A review. *Front. Oncol.* **2023**, *13*, 958310. [\[CrossRef\]](#)
29. Cao, C.-L.; Li, Q.; Tong, J.; Shi, L.; Li, W.-X.; Xu, Y.; Cheng, J.; Du, T.T.; Li, J.; Cui, X.W. Artificial intelligence in thyroid ultrasound. *Front. Oncol.* **2023**, *13*, 1060702.
30. Miranda-Filho, A.; Lortet-Tieulent, J.; Bray, F.; Cao, B.; Franceschi, S.; Vaccarella, S.; Dal Maso, L. Thyroid cancer incidence trends by histology in 25 countries: A population-based study. *Lancet Diabetes Endocrinol.* **2021**, *9*, 225–234. [\[CrossRef\]](#)
31. Locati, L.; Cavalieri, S.; Dal Maso, L.; Busco, S.; Anderson, L.; Botta, L.; Bento, M.J.; Carulla, M.; Chirlaque López, M.D.; Fusco, M.; et al. Rare thyroid malignancies in Europe: Data from the RARECAREnet project. *Oral Oncol.* **2020**, *104*, 104637.
32. Bikas, A.; Burman, K. Epidemiology of thyroid cancer. In *The Thyroid and Its Diseases*; Springer: Cham, Switzerland, 2019; pp. 1–12.
33. Rybakov, S. Medullary thyroid cancer: Epidemiology. *Int. J. Endocrinol.* **2023**, *19*, 306–311. [\[CrossRef\]](#)
34. Mattingly, A.S.; Noel, J.E.; Orloff, L. A closer look at “taller-than-wide” thyroid nodules: Examining dimension ratio to predict malignancy. *Otolaryngol. Head Neck Surg.* **2021**, *165*, 563–570. [\[CrossRef\]](#)
35. Remonti, L.R.; Kramer, C.K.; Leitão, C.B.; Pinto, L.C.; Gross, J.L. Thyroid ultrasound features and risk of carcinoma: A systematic review and meta-analysis of observational studies. *Thyroid* **2015**, *25*, 538–550. [\[CrossRef\]](#)
36. Fukushima, M.; Fukunari, N.; Murakami, T.; Kunii, Y.; Suzuki, S.; Kitaoka, M. Reconfirmation of the accuracy of the taller-than-wide sign in multicenter collaborative research in Japan. *Endocr. J.* **2021**, *68*, 789–797. [\[CrossRef\]](#) [\[PubMed\]](#)
37. Moon, H.; Kwak, J.; Kim, E.K.; Kim, M.J. A taller-than-wide shape in thyroid nodules in transverse and longitudinal planes and the prediction of malignancy. *Thyroid* **2011**, *21*, 1249–1253. [\[CrossRef\]](#) [\[PubMed\]](#)
38. Papapostolou, K.; Evangelopoulou, C.; Ioannidis, I.; Kassi, G.N.; Morfas, K.S.; Karaminas, N.I.; Karga, H.J. Taller-than-wide thyroid nodules with microcalcifications are at high risk of malignancy. *Vivo* **2020**, *34*, 2021–2027. [\[CrossRef\]](#)
39. Zhang, F.; Chen, W. Sonographic features of follicular variant of papillary thyroid carcinoma and performance of the 2017 ACR TI-RADS. *Endocrine* **2020**, *67*, 143–150. [\[CrossRef\]](#) [\[PubMed\]](#)
40. Kim, D.S.; Kim, J.H.; Na, D.G.; Park, S.H.; Kim, E.; Chang, K.; Sohn, C.H.; Choi, Y.H. Sonographic features of follicular variant papillary thyroid carcinomas vs conventional papillary carcinomas. *J. Ultrasound Med.* **2009**, *28*, 1685–1692. [\[CrossRef\]](#)
41. Liu, J.; Zhu, H.; Liang, Z.; Chen, L.; Sun, X.-M.; Shao, Y.; Chen, L. Comparison of the diagnostic performance and clinical role of different ultrasound-based thyroid malignancy risk stratification systems for medullary thyroid carcinoma. *Quant. Imaging Med. Surg.* **2023**, *13*, 3776.

42. Ye, B.B.; Liu, Y.Y.; Zhang, Y.; Liu, B.J.; Guo, L.H.; Wei, Q.; Zhang, Y.F.; Xu, H.X. Predicting tall-cell subtype of papillary thyroid carcinomas independently with preoperative multimodal ultrasound. *Br. J. Radiol.* **2024**, *97*, 1311–1319. [\[CrossRef\]](#)
43. Borowczyk, M.; Woliński, K.; Więckowska, B.; Jodłowska-Siewert, E.; Szczepanek-Parulska, E.; Verburg, F.; Ruchała, M. Sonographic features differentiating follicular thyroid cancer from adenoma—A meta-analysis. *Cancers* **2021**, *13*, 938. [\[CrossRef\]](#)
44. Oh, H.; Kwon, H.; Song, E.; Jeon, M.; Song, D.; Kim, T.; Shong, Y.K.; Baek, J.H.; Kim, W.G. Preoperative predictors for lateral cervical lymph node metastases in sporadic medullary thyroid carcinoma. *Thyroid* **2018**, *28*, 362–368. [\[CrossRef\]](#)
45. Zhang, X.; Huang, W.; Li, X.; Gu, Y.; Jiao, Y.; Dong, F.; Cui, Y. Ultrasound fuzzy entropy imaging based on time-series signal for tissue characterization. *Appl. Acoust.* **2024**, *224*, 110158. [\[CrossRef\]](#)
46. Sollini, M.; Cozzi, L.; Chiti, A.; Kirienko, M. Texture analysis and machine learning to characterize thyroid nodules and differentiated cancer. *Eur. J. Radiol.* **2018**, *99*, 1–8. [\[CrossRef\]](#)
47. Meyer, H.; Schob, S.; Höhn, A.; Surov, A. MRI texture analysis reflects histopathology parameters in thyroid cancer—A preliminary study. *Transl. Oncol.* **2017**, *10*, 911–916. [\[CrossRef\]](#)
48. Nugroho, H.; Rahmawaty, M.; Triyani, Y.; Ardiyanto, I. Texture analysis for classification of thyroid ultrasound images. In Proceedings of the 2016 International Electronics Symposium (IES), Denpasar, Indonesia, 29–30 September 2016; pp. 476–480.
49. Keutgen, X.; Li, H.; Memeh, K.; Busch, J.; Williams, J.; Lan, L.; Sarne, D.; Finnerty, B.; Angelos, P.; Fahey, T.; et al. A machine-learning algorithm for distinguishing malignant from benign indeterminate thyroid nodules using ultrasound radiomic features. *J. Med. Imaging* **2022**, *9*, 034501. [\[CrossRef\]](#) [\[PubMed\]](#)
50. Chiu, L.; Chen, A. A variance-reduction approach to detection of the thyroid nodule boundary on ultrasound images. *Ultrason. Imaging* **2019**, *41*, 206–230. [\[CrossRef\]](#) [\[PubMed\]](#)
51. Chen, J.; Chen, J. Multimodal image feature fusion for improving medical ultrasound image segmentation. *Biomed. Signal Process. Control* **2024**, *89*, 105705. [\[CrossRef\]](#)
52. Ferreira, L.; Gimba, E.; Vinagre, J.; Sobrinho-Simões, M.; Soares, P. Molecular aspects of thyroid calcification. *Int. J. Mol. Sci.* **2020**, *21*, 7718. [\[CrossRef\]](#)
53. Shi, C.; Li, S.; Shi, T.; Liu, B.; Ding, C.; Qin, H. Correlation between thyroid nodule calcification morphology on ultrasound and carcinoma. *J. Int. Med. Res.* **2012**, *40*, 350–357. [\[CrossRef\]](#)
54. Lu, Y.; Zhang, W.; Bai, W.; He, W. Relationship between morphologic characteristics of ultrasonic calcification in thyroid nodules and carcinoma. *Ultrasound Med. Biol.* **2019**, *45*, 2652–2659.
55. Lei, Z.; Li, M.; Luo, D.; Han, Z. The clinical significance of ultrasound grayscale ratio in differentiating markedly hypoechoic and anechoic thyroid nodules. *J. Cancer Res. Ther.* **2018**, *14*, 1567–1571.
56. Nugroho, H.A.; Zulfanahri; Nugroho, A.; Frannita, E.L.; Ardiyanto, I.; Choridah, L. Feature extraction based on laws' texture energy for lesion echogenicity classification of thyroid ultrasound images. In Proceedings of the 2017 International Conference on Computer, Control, Informatics and its Applications (IC3INA), Jakarta, Indonesia, 23–26 October 2017.
57. Sharda, D.N.; Shah, D.A. A study of evaluation of focal thyroid nodule/nodules on ultrasound (grey scale), color doppler and elastography with histopathology correlation. *Int. J. Radiol. Diagn. Imaging* **2023**, *6*, 4–23. [\[CrossRef\]](#)
58. Ghasemi, A.; Zahediasl, S. Normality tests for statistical analysis: A guide for non-statisticians. *Int. J. Endocrinol. Metab.* **2012**, *10*, 486–489. [\[CrossRef\]](#)
59. Dunn, O.J. Multiple comparisons using rank sums. *Technometrics* **1964**, *6*, 241–252. [\[CrossRef\]](#)
60. Šidák, Z. Rectangular confidence regions for the means of multivariate normal distributions. *J. Am. Stat. Assoc.* **1967**, *62*, 626–633. [\[CrossRef\]](#)
61. Lange, T.M.; Gültas, M.; Schmitt, A.; Heinrich, F. optRF: Optimising random forest stability by determining the optimal number of trees. *BMC Bioinform.* **2025**, *26*, 95. [\[CrossRef\]](#)
62. Probst, P.; Boulesteix, A. To tune or not to tune the number of trees in random forest? *J. Mach. Learn. Res.* **2018**, *18*, 1–18.
63. Oshiro, T.M.; Perez, P.S.; Baranauskas, J.A. How many trees in a random forest? In *International Workshop on Machine Learning and Data Mining in Pattern Recognition*; Springer: Berlin/Heidelberg, Germany, 2012.
64. Liu, H.; Haig, E. Semi-random partitioning of data into training and test sets in granular computing. *Granul. Comput.* **2017**, *2*, 357–386. [\[CrossRef\]](#)
65. Chen, R.; Dewi, C.; Huang, S.W.; Caraka, R. Selecting critical features for data classification based on machine learning methods. *J. Big Data* **2020**, *7*, 52. [\[CrossRef\]](#)
66. Iranzad, R.; Liu, X. A review of random forest-based feature selection methods for data science education and applications. *Int. J. Data Sci. Anal.* **2024**, *20*, 197–211. [\[CrossRef\]](#)
67. Zeng, G. On impurity functions in decision trees. *Commun. Stat. Theory Methods* **2024**, *54*, 701–719. [\[CrossRef\]](#)
68. Jaworski, M.; Duda, P.; Rutkowski, L. New splitting criteria for decision trees in stationary data streams. *IEEE Trans. Neural Netw. Learn. Syst.* **2018**, *29*, 2516–2529. [\[CrossRef\]](#)

69. Schoultz, E.; Moccia, C.; Liang, S.; Johansson, E.; Nilsson, M. Tumor Cell Plasticity and Stromal Microenvironment Distinguish Papillary and Follicular Growth Patterns in a Mouse Model of BRAFV600E-induced Thyroid Cancer. *Cancer Res. Commun.* **2025**, *5*, 409–421. [\[CrossRef\]](#)
70. Basolo, F.; Ugolini, C. Pathology of Thyroid Cancer. In *Oxford Textbook of Endocrinology and Diabetes 3e*; Oxford University Press: Oxford, UK, 2021.
71. Williams, D. Thyroid Growth and Cancer. *Eur. Thyroid J.* **2015**, *4*, 164–173. [\[CrossRef\]](#)
72. Suster, S. Thyroid tumors with a follicular growth pattern: Problems in differential diagnosis. *Arch. Pathol. Lab. Med.* **2006**, *130*, 984–988. [\[CrossRef\]](#) [\[PubMed\]](#)
73. Hernandez-Prera, J.; Wenig, B.M. RAS-Mutant Follicular Thyroid Tumors: A Continuous Challenge for Pathologists. *Endocr. Pathol.* **2024**, *35*, 167–184. [\[CrossRef\]](#) [\[PubMed\]](#)
74. Machens, A.; Lorenz, K.; Weber, F.; Dralle, H. Anatomic Patterns of Nodal Spread in Unilateral Papillary and Medullary Thyroid Cancer. *Thyroid* **2024**, *34*, 871–879. [\[CrossRef\]](#) [\[PubMed\]](#)
75. Li, L.; Lian, S.; Luo, Z.; Wang, B.; Li, S. Contour-aware consistency for semi-supervised medical image segmentation. *Biomed. Signal Process. Control* **2024**, *89*, 105694. [\[CrossRef\]](#)
76. Chen, S.; Luo, C.; Liu, S.; Li, H.; Liu, Y.; Zhou, H.; Liu, L.; Chen, H. LD-UNet: A long-distance perceptual model for segmentation of blurred boundaries in medical images. *Comput. Biol. Med.* **2024**, *168*, 107826. [\[CrossRef\]](#)
77. Bi, L.; Fulham, M.; Kim, J. Hyper-fusion network for semi-automatic segmentation of skin lesions. *Med. Image Anal.* **2021**, *72*, 102126. [\[CrossRef\]](#)
78. Cui, W.; Meng, D.; Lu, K.; Wu, Y.; Pan, Z.; Li, X.; Sun, S. Automatic segmentation of ultrasound images using SegNet and local Nakagami distribution fitting model. *Biomed. Signal Process. Control* **2023**, *80*, 104307. [\[CrossRef\]](#)
79. Ma, S.; Li, X.; Tang, J.; Guo, F. Aggregate-aware model with bidirectional edge generation for medical image segmentation. *Appl. Soft Comput.* **2024**, *150*, 111856. [\[CrossRef\]](#)
80. Malhi, H.; Velez, E.; Kazmierski, B.; Gulati, M.; Deurdulian, C.; Cen, S.; Grant, E. Peripheral Thyroid Nodule Calcifications on Sonography: Evaluation of Malignant Potential. *AJR Am. J. Roentgenol.* **2019**, *213*, 672–675. [\[CrossRef\]](#) [\[PubMed\]](#)
81. Nabahati, M.; Ghaemian, N.; Moazezi, Z.; Mehraeen, R. Different sonographic features of peripheral thyroid nodule calcification and risk of malignancy: A prospective observational study. *Pol. J. Radiol.* **2021**, *86*, e366–e371. [\[CrossRef\]](#)
82. Radu, I.; Gheorghe, A.; Sima, O.; Carsote, M.; Nistor, C. Eggshell Calcifications at Thyroid Ultrasound: A Sample-focused Analysis of Cytological Findings and Post-thyroidectomy Pathological Correlates. *Rom. J. Mil. Med.* **2024**, *127*, 428–440.
83. Kim, B.; Choi, Y.; Kwon, H.; Lee, J.; Heo, J.; Han, Y.; Park, Y.; Kim, J. Relationship between patterns of calcification in thyroid nodules and histopathologic findings. *Endocr. J.* **2013**, *60*, 155–160. [\[CrossRef\]](#) [\[PubMed\]](#)
84. Taki, S.; Terahata, S.; Yamashita, R.; Kinuya, K.; Nobata, K.; Kakuda, K.; Kodama, Y.; Yamamoto, I. Thyroid calcifications: Sonographic patterns and incidence of cancer. *Clin. Imaging* **2004**, *28*, 368–371. [\[CrossRef\]](#)
85. Shin, H.; Na, D.; Paik, W.; Yoon, S.; Gwon, H.; Noh, B.; Kim, W. Malignancy Risk Stratification of Thyroid Nodules with Macrocalcification and Rim Calcification Based on Ultrasound Patterns. *Korean J. Radiol.* **2021**, *22*, 663–671. [\[CrossRef\]](#)
86. Ünal, F.; Canpolat, A.; Elhan, A.; Sevim, S.; Sak, S.; Emral, R.; Demir, Ö.; Güllü, S.; Erdoğan, M.; Çorapçıoğlu, D.; et al. Cancer rates and characteristics of thyroid nodules with macrocalcification. *Endocrine* **2024**, *84*, 1021–1029. [\[CrossRef\]](#)
87. Kobaly, K.; Kim, C.; Langer, J.; Mandel, S. Macrocalcifications Do Not Alter Malignancy Risk Within the American Thyroid Association Sonographic Pattern System When Present in Non-High Suspicion Thyroid Nodules. *Thyroid* **2021**, *31*, 1542–1548. [\[CrossRef\]](#)
88. Peng, W.; Qian, Y.; Shi, Y.; Chen, S.; Chen, K.; Xiao, H. Differential Diagnosis of Malignant Thyroid Calcification Nodule Based on Computed Tomography Image Texture. *J. Med. Imaging Health Inform.* **2021**, *11*, 767–772. [\[CrossRef\]](#)
89. Januś, D.; Kujdowicz, M.; Kiszka-Wiłkojć, A.; Kaleta, K.; Taczanowska-Niemczuk, A.; Radliński, J.; Możdżeń, K.; Nowak, Z.; Górecki, W.; Starzyk, J.B. Ultrasound and histopathological assessment of benign, borderline, and malignant thyroid tumors in pediatric patients: An illustrative review and literature overview. *Front. Endocrinol.* **2025**, *15*, 1481804. [\[CrossRef\]](#)
90. Richman, D.M.; Benson, C.; Doubilet, P.; Peters, H.; Huang, S.A.; Asch, E.; Wassner, A.J.; Smith, J.R.; Cherella, C.E.; Frates, M. Thyroid nodules in pediatric patients: Sonographic characteristics and likelihood of cancer. *Radiology* **2018**, *288*, 591–599. [\[CrossRef\]](#)
91. Cappelli, C.; Castellano, M.; Pirola, I.; Gandossi, E.; De Martino, E.; Cumetti, D.; Agosti, B.; Rosei, E. Thyroid nodule shape suggests malignancy. *Eur. J. Endocrinol.* **2006**, *155*, 203–209. [\[CrossRef\]](#) [\[PubMed\]](#)
92. Mohammad, F.; AlZoubi, A.; Du, H.; Jassim, S. Machine learning assessment of border irregularity of thyroid nodules from ultrasound images. In *Multimodal Image Exploitation and Learning 2022*; SPIE: Bellingham, WA, USA, 2022.
93. Batawil, N.; Alkordy, T. Ultrasonographic features associated with malignancy in cytologically indeterminate thyroid nodules. *Eur. J. Surg. Oncol.* **2014**, *40*, 182–186. [\[CrossRef\]](#)
94. Chen, D.; Hu, J.; Zhu, M.; Tang, N.; Yang, Y.; Feng, Y. Diagnosis of thyroid nodules for ultrasonographic characteristics indicative of malignancy using random forest. *BioData Min.* **2020**, *13*, 16. [\[CrossRef\]](#) [\[PubMed\]](#)

95. Cho, K.E.; Gweon, H.M.; Park, A.Y.; Yoo, M.R.; Kim, J.-A.; Youk, J.H.; Park, Y.M.; Son, E.J. Ultrasonographic features of medullary thyroid carcinoma: Do they correlate with pre- and post-operative calcitonin levels. *Asian Pac. J. Cancer Prev.* **2016**, *17*, 3357–3362.
96. Yoon, J.; Kim, E.; Hong, S.; Kwak, J.; Kim, M. Sonographic Features of the Follicular Variant of Papillary Thyroid Carcinoma. *J. Ultrasound Med.* **2008**, *27*, 1431–1437. [[CrossRef](#)]
97. Sillery, J.C.; Reading, C.C.; Charboneau, J.W.; Henrichsen, T.L.; Hay, I.D.; Mandrekar, J.N. Thyroid follicular carcinoma: Sonographic features of 50 cases. *AJR Am. J. Roentgenol.* **2009**, *192*, 795–800. [[CrossRef](#)] [[PubMed](#)]
98. Wu, M.-H.; Chen, C.-N.; Chen, K.Y.; Ho, M.; Tai, H.; Wang, Y.-H.; Chen, A.; Chang, K.J. Quantitative analysis of echogenicity for patients with thyroid. *Sci. Rep.* **2016**, *6*, 35632. [[CrossRef](#)]
99. Zheng, Z.; Liang, E.; Zhang, Y.; Weng, Z.; Chai, J.; Bu, W.; Xu, J.; Su, T. A segmentation-based algorithm for classification of benign and malignancy thyroid nodules with multi-feature information. *Biomed. Eng. Lett.* **2024**, *14*, 785–800. [[CrossRef](#)] [[PubMed](#)]
100. Aboudi, N.; Guetari, R.; Khelifa, N. Multi-objectives optimisation of features selection for the classification of thyroid nodules in ultrasound images. *IET Image Process.* **2020**, *14*, 1901–1908. [[CrossRef](#)]
101. Acharya, U.; Vinitha Sree, S.; Krishnan, M.M.R.; Molinari, F.; Garberoglio, R.; Suri, J.S. Non-invasive automated 3D thyroid lesion classification in ultrasound: A class of ThyroScan™ systems. *Ultrasonics* **2012**, *52*, 508–520. [[CrossRef](#)]
102. Mugasa, H.; Dua, S.; Koh, J.E.W.; Hagiwara, Y.; Oh, S.L.; Madla, C.; Kongmebol, P.; Ng, K.; Acharya, U.R. An adaptive feature extraction model for classification of thyroid lesions in ultrasound images. *Pattern Recognit. Lett.* **2020**, *131*, 463–473. [[CrossRef](#)]
103. Song, R.; Zhang, L.; Zhu, C.; Liu, J.; Yang, J.; Zhang, T. Thyroid Nodule Ultrasound Image Classification Through Hybrid Feature Cropping Network. *IEEE Access* **2020**, *8*, 64064–64074.
104. Xing, G.; Miao, Z.; Zheng, Y.; Zhao, M. A multi-task model for reliable classification of thyroid nodules in ultrasound images. *Biomed. Eng. Lett.* **2023**, *14*, 187–197. [[CrossRef](#)]
105. Li, P.; Tao, H.; Zhou, H.; Zhao, M. Enhanced Multiview attention network with random interpolation resize for few-shot surface defect detection. *Multimed. Syst.* **2025**, *31*, 36. [[CrossRef](#)]
106. Wang, Z.; Tao, H.; Zhou, H.; Deng, Y.; Zhou, P. A content-style control network with style contrastive learning for underwater image enhancement. *Multimed. Syst.* **2025**, *31*, 60.
107. Apedo, Y.; Tao, H. A weakly supervised pavement crack segmentation based on adversarial learning and transformers. *Multimed. Syst.* **2025**, *31*, 266.
108. Chi, J.; Walia, E.; Babyn, P.; Wang, J.; Groot, G.; Eramian, M. Thyroid Nodule Classification in Ultrasound Images by Fine-Tuning Deep Convolutional Neural Network. *J. Digil. Imaging* **2017**, *30*, 477–486.
109. Wang, Y.; Yue, W.; Li, X.; Liu, S.; Guo, L.; Xu, H.; Zhang, H.; Yang, G. Comparison Study of Radiomics and Deep Learning-Based Methods for Thyroid Nodules Classification Using Ultrasound Images. *IEEE Access* **2020**, *8*, 52010–52017.
110. Horvath, E.; Silva, C.; Majlis, S.; Rodriguez, I.; Skoknic, V.; Castro, A.; Rojas, H.; Niedmann, J.; Madrid, A.; Capdeville, F.; et al. Prospective validation of the ultrasound based TIRADS (Thyroid Imaging Reporting and Data System) classification: Results in surgically resected thyroid nodules. *Eur. Radiol.* **2017**, *27*, 2619–2628. [[PubMed](#)]

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.